Data Management Plan
WP6 - Management, Coordination and Dissemination

Deliverable Title	Data Management Plan
Deliverable number	D6.3
Brief Description	Data Management Plan describing how the data collected during the research will be handed during and after the end of the project (updated version in M36)
WP number	WP6
Lead Beneficiary	FONDAZIONE CMCC
Author(s)	Emily Tynes (AU); Fabiana Gerardi, Soheil Shayegh (CMCC); Luke Kirwan, Matthew Gidden, Philip Hackstock (IIASA), Vassilis Daioglou (PBL)
Deliverable Due Date	29/02/2024
Actual Delivery Date	29/02/2024
Nature of the Deliverable	DMP ― Data Management Plan
Dissemination Level	PU - Public
KeyWords	Datasets, Metadata, FAIR data, Repository, Open access, Data security and Ethics

Date	Ver.	Contributors	Pages

GLOSSARY

Acronim	Full Terms
CDR	Carbon Dioxide Removal
DMP	Data Management Plan
DOI	Digital Object Identifier
FAIR	Findable, Accessible, Interoperable, and Re-usable
IPR	Intellectual Property Rights
SSH	Social Sciences and Humanities
WP D T	Work package Deliverable Task

LIST OF CONTENTS

1. Executive Summary

The UPTAKE Data Management Plan (DMP) addresses the **types and nature** of data collected and/or generated during the project lifespan. It summarizes the critical data and how it will be managed to ensure compliance with the **FAIR principles** – Findable, Accessible, Interoperable, and Re-usable.

The DMP will also address the **open science practices** that the UPTAKE project will implement for any new model or tool developed or substantially improved within the project and how the Consortium will ensure documentation, availability of model code, and input data developed within the project.

A **revised version of the DMP** will be submitted at the later stages of the project based on this deliverable and reflecting the progress made in data production throughout the project's lifespan (Month 36 - August 2026).

2. Data Summary

The UPTAKE project combines **various methods and tools** due to the unique variety of researchers involved in the action (from engineering to climate-economic modelling to political science and Social Sciences and Humanities SSH assessment).

Figure 1 will help to summarize the whole methodological approach in the Project and help to identify the three pillars of the project: **improved Carbon Dioxide removal CDR knowledge**, **enhanced systematic CDR assessment**, and **encompassing CDR enabling frameworks**, which figure out the main output datasets foreseen by the project:

1. **CDR Stakeholder Forum** (D1.2; D1.7;D6.5)
2. **CDR Knowledge Inventory** (D5.1)
3. **CDR Roadmap Explorer** (D5.2 and D5.5)
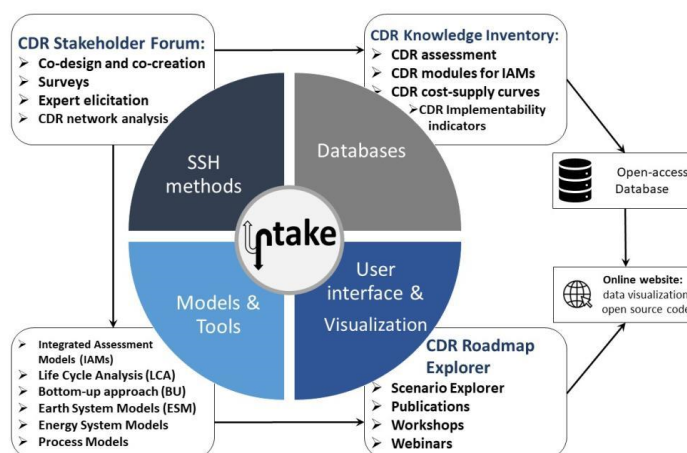
The datasets of the UPTAKE project will combine existing open-source data and knowledge about **CDR technologies** (CDR Knowledge Inventory), mitigation pathways, and **CDR uptake scenarios** (CDR Roadmap Explorer), with data related to the CDR Stakeholder forum and the **surveys** carried out in WP1 (CDR Stakeholder Forum).

The survey undertakes at least six nationally representative surveys or focus groups to measure public attitudes and expectations of the same countries identified in T1.2. A set of experiments would help to investigate what CDR technologies more easily lend themselves to misunderstanding, the generation of fake information, and how to convey information best to minimize these risks. Survey data will be anonymized, and no sensitive information (including contact details or names) will be elicited from the respondents. he data collected from surveys in T1.3 will be stored in a secure, password-encrypted server maintained by Aarhus University (AU) IT. The Ethics Management Plan (D6.4) will assess detailed information on survey data management.

2.1 New data generation

Data visualisation and a CDR Scenario Explorer

A new IIASA **Scenario Explorer** will be developed to host **UPTAKE data** in a public, transparent platform that enables user-friendly exploration of the scientific outcomes and findings. The Explorer infrastructure will be enhanced and further developed specifically for unique assessments to be undertaken within the project, including support for spatial maps of CDR potentials (WP2), open access to the regional and country-level results of CDR strategies (WP3) and the inventory of implementability indicators for CDR uptake (WP4).

The Scenario Explorer is currently the most widely used **data hub** for integrated assessment modelling comparisons. It supports the IPCC WG3 scenario assessments and is also widely used in several projects with a global (e.g., H2020 projects CD-LINKS, ENGAGE, NAVIGATE, ELEVATE) and European focus (e.g., SET-Nav, ECEMF, ARIADNE and openENTRANCE). It serves as a central data hub for modellers, **facilitating cooperation and exchange across countries and modelling groups**. The main feature is an intuitive web user interface that enables experts and a wider audience to visualize and compare scenario results and input data. The infrastructure also supports automated validation of submitted scenarios, allowing retrieving data via a public API. Snapshots of data produced by UPTAKE will be deposited on Zenodo, have a DOI assigned and comply with the FAIR principles.

Stakeholder Forum

The online **Stakeholder Platform** will support the open and interactive CDR Stakeholder Forum, which will be used to ensure consistent and **efficient involvement of stakeholders** in the activities implemented within (and beyond) the Project; **disseminate the interim results** using the interactive tools from Data visualisation and a CDR Scenario Explorer among the key stakeholders and **receive feedback** to improve the project outcomes; stimulate constructive, evidence-based dialogue among different stakeholder groups from scientific, policymaking and business communities as well as social groups active in the CDR debate and impacted communities.

The Forum, managed by REFORM in close cooperation with CMCC, will provide impulses for the exchanges involving stakeholders and the project team throughout the project co-design and co-creation embedded throughout the Project. All data from the UPTAKE Stakeholder Forum will be securely stored in the CMCC Sever.

CDR Knowledge inventory

We will produce a **synthesis report** of the CDR knowledge, including data and a description of the new CDR deployment narratives. This knowledge inventory will synthesise knowledge generated from all the WPs, combining techno-economic inventories, implementability and feasibility aspects of CDR, and estimates of realistic and implementable potentials of CDR. As such, this inventory will be an important resource for understanding and interpreting the Resilient and Implementable CDR roadmaps of Task 5.4. This inventory will be **made public as a report**, and a **dedicated website** will host a database of the gathered information. It is intended to give a snapshot of the database a DOI and make it compliant with FAIR principles. This knowledge inventory will provide a resource to be used by the scientific and policy community within and beyond the UPTAKE project.

2.2 Existing data

The Consortium agreed to share data, models and software developed in other collaborative projects, which has been indicated as relevant to implementing UPTAKE actions. The Consortium Agreement (CA) Attachment 1: Background includes all conditions and specific limitations to the use of this existing Data/Model/Software.

Apart from important models needed for implementing the action (such as: CMCC's WITCH model; PIK's REMIND model; MCC's PRIDE model; etc[1]) IIASA will provide existing climate mitigation scenario data developed at the institute, as well as underlying modeling methodologies.

The climate mitigation scenario data is produced by the open-access model MESSAGEix. Some of the data are collected in databases with specific licenses :

(https://data.ene.iiasa.ac.at/ngfs/#/license or

https://data.ene.iiasa.ac.at/engage/#/license)

Furthermore, ecoinvent is the Life Cycle Assessment (LCA) background database, which will be used as source of background life **cycle inventory data** within the UPTAKE project (PSI).

---

[1] *According to the Grant Agreement (Article 16.1) Background is defined as "data, know-how or information (…) that is (…) needed to implement the Action or exploit the results". Because of this need, access rights have to be granted in principle, but parties must identify and agree amongst them on the Background for the project. This is the purpose of the Attachment 1, included in the UPTAKE Consortium Agreement.*

3. FAIR data

3.1 Making data findable, including provisions for metadata

The UPTAKE consortium is dedicated to contributing to **Open Science**, in line with the EU recommendations on **Open Access**, **Open Source** and **Open Data**.

Following the examples of the established open-source models (e.g., WITCH, REMIND, MESSAGEix) and scientific-software packages (e.g., Pyam) of consortium partners, new tools, datasets and modules developed in the UPTAKE project will be deposited in the open-source repository: ZENODO (or suitable alternatives) and given a unique identifier (such as a DOI).

3.2 Making data accessible

The data generated by the UPTAKE project will be made freely available and user-friendly principally through:

- The **Scenario Explorer** hosted by IIASA
- The dedicated community on the **online repository** ZENODO

Furthermore, the Consortium will follow an open-source model, i.e., the results of the projects will be licensed under a commonly accepted license agreement such as Creative Commons or Apache.

For all peer-reviewed scientific publications related to the results of UPTAKE project research, open access (free of charge online access for any user) will be ensured. We aim to publish high-quality, primarily **Open Access journals** (Gold Open Access).

Preprints of articles in non-open access journals will be made available through institutional repositories or the Open Access Infrastructure for Research in Europe (OpenAIRE). All partners have been allocated additional funding to cover some publishing costs for Gold Open Access to ensure free access to the research outputs.

3.3 Making data interoperable

UPTAKE members already have strong track records in open-source software and models. Besides the public release of the software, UPTAKE aims to **make models transparent and accessible** by carefully documenting them and by **associating models with high-quality metadata** with clear identifiers of model inputs and outputs. Data and scenarios will be appropriately indexed to make them **easily searchable and retrievable**. Appropriate software

and programming languages will be used **to ensure the interoperability of data and models**. For scenario outcomes, UPTAKE will leverage pyam, https://pyam-iamc.readthedocs.io/en/stable/, an open-source Python package developed by the IAM community from this project. Pyam facilitates the reproducibility and reliability of scenario processing, validation, and analysis. The modules and datasets developed in UPTAKE will be oriented on these conventions to **ensure maximum Interoperability and Reusability of data/research outputs**.

3.4 Increase data re-use

Datasets produced in UPTAKE are also intended to be used outside of the project. Using open-source repositories such as ZENODO will help anyone to use, change, and share. Furthermore, the IIASA Scenario Explorer for UPTAKE will follow standard FAIR practices, including providing data with an open-access license, documenting the standard IAM data (e.g., model, scenario, region, and variable definitions) on a documentation page, and providing easy-entry user-facing features such as dashboards and workspaces. All data housed in the IIASA Scenario Explorer will be tagged with a DOI and co-hosted on ZENODO.

4. Other research outputs

The UPTAKE project is not expecting to produce any other outputs.

5. Allocation of resources

The budget assigned to each party of the UPTAKE project includes costs for making the data generated by the project, FAIR. Research partners within the project will assume the costs of open access, where necessary, and will post versions of published manuscripts in self-archived online repositories as needed.

6. Data security

The project's data storage will be mainly curated by AU and CMCC. In particular, survey data will be collected and maintained by the GDPR-compliant survey provider, with which AU has an agreement through the AU Technology Transfer Office. They potentially contain data such as name, contact information, gender, racial and ethnic origin, and political and philosophical opinions, defined as identifiable sensitive personal data by GDPR. Thus, the data will be

transmitted and stored in a way that protects it from unauthorised access. Data from survey providers will be transmitted in pseudonymised form and stored on **password-encrypted servers** as part of the AU network. Pseudonymised data will only be shared amongst Consortium partners that appear on the informed consent letter. Once the survey has been completed and all data has been collected, the provider will delete any remaining identifiable information within three months, making the data fully anonymised. At that time, the data can be shared amongst the entire Consortium. CMCC will take care of the Stakeholder Forum data, which will be stored on a password-encrypted server maintained by CMCC. The Scenario Explorer data will be stored by IIASA instead: snapshots of data produced by UPTAKE will be deposited on ZENODO, assigned a DOI, and comply with the FAIR principles (see section 2.1 of this document).

7. Ethics

The research proposed in UPTAKE involves collecting data through the participation of adult and healthy volunteers. In particular, the Project will collect data during the **stakeholder engagement** and **vision development process**; it will conduct **surveys** on public attitudes towards CDR and interviews with CDR stakeholders; it will collect personal data during the registration process for the **webinar series**. The activities foreseen do not involve children, minors, or persons with disabilities.

Data gathering, analysis, and storage for research at all organizations involved in the Project will be conducted following the General Data Protection Regulation (EU) 2016/679. All ethical aspects of the UPTAKE project, both management and data assessment, are defined in deliverable 6.4, _Ethics Management Plan_. Please refer to that document for a deeper description.

8. Other issues

The UPTAKE data management is granted by **Article 16** of the UPTAKE **Grant Agreement** (GA) and **Article 9** of the **CA**, which also refers to the detailed Background identified and agreed upon by the Consortium for the research activities carried out in the Project.
The only other reference to different procedures and use of data in the UPTAKE project refers to the access and use of **_ecoinvent data and Database_** listed in the Background of the Swiss-

associated partner, Paul Scherrer Institut (PSI). The use of ecoinvent within the UPTAKE project must comply with the **End User Licence Agreement** (EULA), a contract between the ecoinvent Association and any individuals or legal entities using the ecoinvent Database. Its terms of use apply specifically to the use of data and databases and are governed by Swiss law. Complete Terms of use of ecoinvent EULA is accessible through https://ecoinvent.org/the-ecoinvent-database/terms-of-use/.